

Construir un prototipo de software como ayuda para la detección del cáncer de próstata a través de análisis de imágenes por resonancia magnética

Cristian Camilo Holguín Gutiérrez

John Edward Narváez Londoño

Universidad Tecnológica de Pereira

Facultad de ingenierías

Ingeniería de sistemas y computación

Construir un prototipo de software como ayuda para la detección del cáncer de próstata a través de análisis de imágenes por resonancia magnética

PhD. Guillermo Roberto Solarte

Director proyecto de grado

Universidad Tecnológica de Pereira

Facultad de ingenierías

Ingeniería de sistemas y computación

Contenido

Dedicatoria	5
Introducción	6
Descripción del problema	6
Justificación	6
Objetivos	7
• Objetivo general	7
• Objetivos específicos	7
Hipótesis	7
Alcance y limitaciones	7
Marco teórico	8
Marco referencial	8
• Estado del arte	8
• Marco conceptual	9
• Marco legal	9
Metodología	9
Análisis arquitectónico del prototipo usando el método de análisis de compensación de arquitectura (ATAM)	11
• Drivers arquitectónicos	11
• Lista de escenarios	11
• Especificación de escenarios	13
• Árbol de utilidad	15
Toma de requerimientos	16
Requisitos funcionales	16
Requisitos no funcionales	16
Inversos	16
Diagrama de casos de uso	17
Diagramas de estados	18
Diagramas de actividades	21
Diagramas de secuencias	24
Diagramas de paquetes	27
Desarrollo y resultados	28

Conclusiones	35
Aportes	36
Recomendaciones	36
Bibliografía	37

Lista de tablas

Tabla 1: Drivers, fuente: propia	12
Tabla 2: Lista de escenarios, fuente: propia	13
Tabla 3: Especificación escenario 1, fuente: propia	13
Tabla 4: Especificaciones escenario 2, fuente: propia	13
Tabla 5: Especificaciones escenario 3, fuente: propia	13
Tabla 6: Especificaciones escenario 4, fuente: propia	14
Tabla 7: Especificación de escenario 5, fuente: propia	14
Tabla 8: Especificación escenario 6, fuente: propia	14
Tabla 9: Especificación escenario 7, fuente: propia	14

Lista de ilustraciones

Ilustración 1: Árbol de utilidad, fuente: propia	15
Ilustración 2: Casos de uso, fuente: propia	17
Ilustración 3: DE-Registro de usuarios, fuente: propia	18
Ilustración 4: DE-Ingresa usuario, fuente: propia	19
Ilustración 5: DE-Ingresa contenido, fuente: propia	20
Ilustración 6: DA-Ingresa usuario, fuente: propia	21
Ilustración 7: DA-Registrar usuario, fuente: propia	22
Ilustración 8: DA-Ingresa contenido, fuente: propia	23
Ilustración 9: DS-Ingresa usuario, fuente: propia	24
Ilustración 10: DS-Registrar usuario, fuente: propia	25
Ilustración 11: DS-Ingresa contenido, fuente: propia	26
Ilustración 12: Diagrama de paquetes, fuente: propia	27
Ilustración 13: Imagen de resonancia magnética, toma axial, fuente: Cancer imaging archive	29
Ilustración 14: Imagen de resonancia magnética, toma coronal, fuente: Cancer imaging archive	29
Ilustración 15: Imagen de resonancia magnética, toma sagital, fuente: Cancer imaging archive	30
Ilustración 16: Vista de inicio de sesión, fuente: propia	31
Ilustración 17: Vista ingreso datos, fuente: propia	31

Ilustración 18: Vista resultado evaluación, fuente: propia	32
Ilustración 19: Parámetros para la red neuronal, fuente: Propia.....	32
Ilustración 20: Esquema red neuronal convolucional, fuente: https://www.diegocalvo.es/ ..	33
Ilustración 21: Capas de inicio, convolución y agrupación, fuente: propia	33
Ilustración 22: Capa aplanar, fuente: propia.....	33
Ilustración 23: Capa densa, fuente: propia	34
Ilustración 24: Capas de expulsión, fuente: propia.....	34
Ilustración 25: Capa densa para elección de clase, fuente: propia	34
Ilustración 26: Capas de salida, fuente: propia.....	34
Ilustración 27: Evidencia de red entrenada, fuente: propia	35
Ilustración 28: Fragmento función de salida del programa, fuente: propia.....	35

Dedicatoria

Inicialmente agradecer a mi madre quien me ha apoyado incondicionalmente en cada paso hacia mi vida privada y profesional, regalándome esa fe y confianza en mí, que tal vez desde mi perspectiva no pude ver en este trayecto, a mi padrastro y hermanos que han sido fundamentales para mi equilibrio emocional y social para lograr este gran paso.

A los compañeros y profesores que han apoyado todo este viaje universitario.

Cristian Camilo Holguín.

A mis padres por su apoyo incondicional, como primeros maestros de vida y máxima inspiración.

A la Universidad Tecnológica de Pereira por ser nuestro alma máter y formarnos con ciencia y técnica para el beneficio de la humanidad.

A la facultad de Ingeniería de sistemas y computación, profesores y compañeros por su acompañamiento y ética tenaz.

A todos aquellos testigos que nos motivaron a alcanzar este logro académico.

John Edward Narváez

Introducción

Con el uso de las herramientas tecnológicas se busca optimizar un sinnúmero de tareas, entre ellas, está el diagnóstico temprano de diferentes tipos de cáncer. En los últimos años, se han desarrollado diversos tipos de software que le permiten al profesional de la salud apoyarse en la tecnología para lograr la mayor precisión en el diagnóstico de la enfermedad y así definir de manera más ágil el tratamiento adecuado.

Entre las herramientas de diagnóstico podemos mencionar las pruebas clínicas que se basan en biomarcadores extraídos de muestras de fluidos corporales como: sangre, orina, semen o tejido.

Descripción del problema

La población con mayor índice de vulnerabilidad de padecer esta enfermedad tiene poca accesibilidad a la información siendo uno de los problemas más recurrentes, según un artículo del periódico El Espectador en 2017, se tenía información de entre 8000 y 9000 casos nuevos en Colombia, entre los factores de riesgo asociados encontramos la historia clínica familiar, la geografía, raza o grupo étnico y la edad.^{[4][5]}

Justificación

El procedimiento inicial de la detección de la enfermedad ejerce una presión social y/o psicológica en el género masculino, reduciendo así sus probabilidades de tratarse a tiempo y de manera adecuada, además la carente prevención en salud eleva las probabilidades de un diagnóstico fatal, sumado a esto intervienen otras consecuencias de salud como problemas urinarios, disfunción eréctil, hemorroides dolorosas y problemas de tipo intestinal. En la actualidad, el cáncer de próstata se ubica entre las 5 primeras causas de mortalidad en hombres, representando el 15% de ellas en edades entre los 60 y 75 años, muchas de ellas debido al tratamiento inoportuno o nulo de dicha enfermedad.^[4]

Si bien se han realizado investigaciones sobre el tema, por medio de la tecnología se pretende hallar soluciones que resulten más eficaces al momento de diagnosticar la enfermedad y le permitan al profesional realizar con más eficiencia su trabajo. Hablando puntualmente acerca del software y al ser utilizado como una herramienta de apoyo para detectar con mayor precisión las posibles inconsistencias en la zona prostática, se busca generar una nueva alternativa que reduzca la necesidad de practicar el tacto rectal para así reducir la presión social.

Objetivos

- **Objetivo general**

Construir un prototipo de software como ayuda al profesional de la salud para la detección del cáncer de próstata a través de análisis de imágenes obtenidas por resonancia magnética.

- **Objetivos específicos**

- Recopilar algunas investigaciones científicas (artículos) en las cuales se pueda evidenciar síntomas, causas, consecuencias, índice de mortalidad, edad, raza/etnia.
- Diseñar una red neuronal convolucional que permita clasificar las imágenes por resonancia magnética.
- Implementar una red neuronal convolucional para cada corte (axiales, coronales y sagitales).
- Implementar una función que tome los resultados de cada red y los simplifique en una única respuesta.

Hipótesis

Es posible realizar la clasificación positiva o negativa a cáncer de próstata, utilizando redes neuronales convolucionales.

Alcance y limitaciones

El alcance de este proyecto se enfoca en un prototipo cuyo propósito es la exploración diagnóstica del cáncer de próstata por medio de imágenes de resonancia magnética, se evidencia una limitación al acceso de bases de datos que contengan imágenes por resonancia magnética de personas que no padezcan cáncer de próstata.

Marco teórico

Una de las grandes referencias que se tienen a día de hoy con respecto a la integración de la inteligencia artificial en el campo de la medicina a través de algoritmos de deep learning para el reconocimiento de imágenes, es el proyecto iFIND desarrollado en el laboratorio BioMedIA^[13] del Imperial College en Londres, el cual implementó dicha técnica para estudiar las anomalías que se pueden presentar en las fases que atraviesa un feto humano con excelentes resultados. También en el campo oncológico, bajo la premisa de diagnosticar a tiempo el cáncer de mama, la Universidad de Valencia mediante un grupo de expertos desarrolla algoritmos inteligentes para detectar el cáncer mediante el análisis de mamografías con una confiabilidad aproximada del 89% y con el aval de un radiólogo, este porcentaje incrementa un 93%.

Es así como haciendo uso de este método se pretenden procesar imágenes de tres tipos (sagital, axial y coronal) mediante una red convolucional. La red se encarga de evidenciar la presencia de posibles anomalías validadas por un profesional en radiología, que aumentará la fiabilidad, similar al estudio que adelanta la Universidad de Valencia con resultados satisfactorios. De acuerdo con las estimaciones de GLOBOCAN^[14] (2012) se presume, que el riesgo acumulado de tener cáncer de próstata antes de los 75 años es de 4.3%. Además, se considera que uno de cada tres hombres mayores de cincuenta años ha sido diagnosticado con cáncer y más del 60% de los casos nuevos se dan en hombres entre los 60 y 79 años. Es por esto que esta franja en particular se define como la población demográfica objetivo, en cuanto al aspecto geográfico se tienen datos sobre las regiones con más incidencia en este tipo de cáncer según el grupo de vigilancia epidemiológica^[15] INC-cáncer en cifras son: San Andrés y Providencia, Atlántico, Valle del Cauca, Bogotá, Bolívar y Chocó, poblaciones apremiantes y objetivos.

Marco referencial

● Estado del arte

Hasta el momento de la creación de este documento, se encuentran desarrollos clínicos con la finalidad de encontrar nuevas perspectivas en cuanto al diagnóstico de cáncer de próstata, algunos de estos avances desde el punto computacional son los siguientes:

- **Aurora:** Software que consulta bases de datos, le proporciona información que le define el diagnóstico y hallar un tratamiento específico para cada paciente.^[1]
- **4k score:** Prueba de sangre que combina el resultado de las proteínas PSA total, PSA libre, PSA intacto y calicreína 2 humana (hK2); este examen busca reducir significativamente la cantidad de biopsias.^[2]
- **CancerLocator:** Software que puede detectar simultáneamente el cáncer e identificar en qué parte del cuerpo se encuentra mediante la muestra de sangre de un paciente.^[3]

- **Marco conceptual**

Imagen por resonancia magnética (IRM o MRI):

Técnica no invasiva, que permite obtener información sobre la estructura o el cuerpo que se va a analizar, el cual retorna una serie de imágenes que el profesional encargado podrá interpretar.

Cáncer de próstata

Se desarrolla en uno de los órganos glandulares del sistema reproductor masculino, la próstata. Esto se produce cuando hay un crecimiento sin control de algunas células prostáticas que mutan. Existe el riesgo de que este se pueda propagar a otros sitios del cuerpo.

Red neuronal convolucional

Es un tipo de red neuronal artificial que permite procesar imágenes más eficientemente que una tradicional, debido a que su aplicación se realiza en matrices bidimensionales.

- **Marco legal**

- Ley 23 de 1982, Derechos de autor.
- Ley 44 de 1993, Software sin licencia.

Metodología

El proyecto de aplicación se desarrollará utilizando la metodología ágil XP (Xtreme Programing) que elegimos porque el grupo de trabajo es muy reducido, estableciendo una comunicación más simple entre los miembros.

La metodología XP consta de cuatro las fases:

Primera fase (Planificación del proyecto):

Conocer cuáles son las anomalías que podamos encontrar en una imagen por resonancia magnética, además de que investigaciones existen en la actualidad para así limitar el alcance del proyecto.

Para ello:

- **Estado del arte:** Se investigará en diferentes fuentes los posibles factores de riesgo, cómo se diagnostica el cáncer de próstata, las probables consecuencias y lo que se está haciendo en el ámbito de la ciencia computacional para combatir este problema.

- **Análisis de datos:** Con el propósito de conocer la viabilidad del proyecto una vez hecha la investigación, se concertará una reunión para clasificar la información recopilada.
- **Recolección de recursos:** Se debe nutrir la denominada base del conocimiento, para ello tomaremos en cuenta bases de datos de dominio público donde podamos extraer las imágenes por resonancia magnética de la próstata, las cuales serán el recurso vital para el entrenamiento de las redes neuronales.

Segunda fase (Diseño):

De manera inicial usamos el método de análisis de acuerdos de arquitectura (ATAM), para identificar las funciones y necesidades críticas, para así mitigar un posible error desde el comienzo del desarrollo del sistema, además de, tempranamente, tomar mejores decisiones en el proceso de desarrollo.

A partir de los requerimientos tomadas en la fase anterior, iniciamos el diseño del proyecto de forma simple, basándonos en el modelo UML, entonces:

- Elaborar un documento que permita entender el procedimiento que se llevará a cabo.
- Elaborar los diagramas del modelo UML.

Tercera fase (Codificación):

- Se desarrollará una interfaz amigable con el usuario.
- Se desarrollará una base de datos en la cual se almacenarán las imágenes que serán usadas para el entrenamiento de la red neuronal.
- Se implementará una red neuronal convolucional por cada corte, para ello se hará un pre-procesamiento que genera imágenes que el algoritmo pueda usar.

Cuarta fase (Pruebas):

- Pruebas unitarias del sistema que cumplan con los criterios que se hayan establecido.
- Pruebas de integración que nos permiten visualizar su correcto funcionamiento entre módulos.
- Se llevará a cabo una sustentación que nos permita mostrar resultados del prototipo.

Análisis arquitectónico del prototipo usando el método de análisis de compensación de arquitectura (ATAM)

Es una arquitectura que permite mitigar el riesgo en los procesos de manera temprana en el ciclo de vida del desarrollo del software, determina qué casos de uso son de mayor importancia y que posibles riesgos pueden presentar, además de incrementar la comunicación con los interesados.

- **Drivers arquitectónicos**

TIPO DE DRIVER (AC/CU/REST) * AC: Atributo de calidad, CU: Caso de uso, Rest: Restricción	PRIORIDAD	DESCRIPCIÓN DEL DRIVER
1. CU	5	Ingresar contenido a evaluar
2. CU	1	Registrar usuario
3. AC (Escalabilidad)	3	Control de usuarios
4. AC (Modificabilidad)	3	El sistema podrá ser traducido.
5. REST	4	Al sistema se podrá ingresar máximo 1 imagen por cada corte.
6. AC (Facilidad de pruebas)	3	Se podrán hacer pruebas independientes de los módulos.
7. AC (Desempeño)	3	Tiempo estimado de espera máximo para resultados: 10 segundos
8. REST	3	Para un resultado eficiente en el entrenamiento se recomienda un computador que posea procesador de 8 hilos y 8 Gb de RAM

Tabla 1: Drivers, fuente: propia

- **Lista de escenarios**

ID	ESCENARIOS	DRIVER RELACIONA DOS	FUSIONADO	PRIORIDAD
1	El sistema debe validar los datos de ingreso de los usuarios (Usuario y contraseña).	Seguridad		Media
2	El sistema debe notificar cuando los datos ingresados son insuficientes.	Desempeño		Alta
3	El sistema debe notificar cuando los datos de ingreso son incorrectos.	Seguridad		Alta
4	El sistema permitirá elegir el tipo de corte a ingresar.	Desempeño		Media
5	El sistema permitirá hacer análisis a varias entradas de manera simultánea.	Desempeño		Alto
6	El sistema debe permitir el registro de usuarios	Usabilidad		Alto
7	El sistema entregará la información al usuario en una vista con los datos correspondientes del análisis	Usabilidad		Alto

Tabla 2: Lista de escenarios, fuente: propia

- **Especificación de escenarios**

Escenario crudo	1
Estímulo	Intento de ingreso por parte del usuario
Fuente de estímulo	Usuario final
Entorno	Instante normal de operación
Artefacto	Computador
Respuesta	El sistema consulta la existencia del usuario y permite el ingreso o no del usuario
Medida de la respuesta	Menor a 5 segundos
Preguntas	
Problemas	

Tabla 3: Especificación escenario 1, fuente: propia

Escenario crudo	2
Estímulo	Ejecución de función entrega de datos
Fuente de estímulo	Sistema
Entorno	Instante normal de operación
Artefacto	Computador
Respuesta	Corto texto que describe el resultado de la ejecución de la función
Medida de la respuesta	Menor a 10 segundos
Preguntas	
Problemas	

Tabla 4: Especificaciones escenario 2, fuente: propia

Escenario crudo	3
Estímulo	Error en la validación de datos de usuario
Fuente de estímulo	Base de datos
Entorno	Validación de datos
Artefacto	Base de datos
Respuesta	Corto texto que describe el resultado de la ejecución de la función
Medida de la respuesta	Menor a 5 segundos
Preguntas	
Problemas	

Tabla 5: Especificaciones escenario 3, fuente: propia

Escenario crudo	4
Estímulo	Función unitaria de ingreso de datos
Fuente de estímulo	Usuario final
Entorno	Instante normal de operación
Artefacto	Computador
Respuesta	Nueva ventana que permite la selección de imágenes
Medida de la respuesta	Menor a 4 segundos
Preguntas	
Problemas	

Tabla 6: Especificaciones escenario 4, fuente: propia

Escenario crudo	5
Estímulo	Ingreso de varias imágenes
Fuente de estímulo	Usuario final
Entorno	Instante normal de operación
Artefacto	Computador
Respuesta	Nueva vista
Medida de la respuesta	Menor a 10 segundos
Preguntas	
Problemas	

Tabla 7: Especificación de escenario 5, fuente: propia

Escenario crudo	6
Estímulo	Intento de registro por parte del root
Fuente de estímulo	Usuario root
Entorno	Instante normal de operación
Artefacto	Computador
Respuesta	Vista con formulario para crear usuario final
Medida de la respuesta	Menor a 5 segundos
Preguntas	
Problemas	

Tabla 8: Especificación escenario 6, fuente: propia

Escenario crudo	7
Estímulo	Aceptación de imágenes ingresadas
Fuente de estímulo	Usuario final
Entorno	Instante normal de operación
Artefacto	Sistema
Respuesta	Vista que refleja el resultado de la operación
Medida de la respuesta	Menor a 10 segundos
Preguntas	
Problemas	

Tabla 9: Especificación escenario 7, fuente: propia

[illegible]

15

El árbol de utilidad nos presenta los atributos de calidad del software hasta el establecimiento de escenarios que especifican con suficiente detalle el nivel de prioridad de cada uno de ellos. Cada atributo de calidad pertenece al árbol contiene una serie de escenarios relacionados y una escala jerárquica de importancia asociada.

Toma de requerimientos

El prototipo estará en la capacidad de realizar las siguientes funcionalidades:

- Entrenar la red neuronal convolucional para que esté en la capacidad de clasificar ciertos patrones diferenciales de cada corte propuesto.
- Almacenar la base de datos de imágenes (las cuales sean de dominio público) con la cual se entrenará la red neuronal convolucional.
- En base a la evaluación de patrones, puede arrojar un resultado dada una nueva imagen suministrada por el usuario.

Requisitos funcionales

- **Root**
 - Podrá crear un perfil de usuario.
 - Podrá borrar un perfil de usuario.
- **Médico**
 - Podrá agregar imágenes a evaluar en el sistema.
 - Podrá eliminar imágenes suministradas al sistema.
 - Podrá editar datos del usuario paciente.
- **Paciente**
 - Solo mostrará información registrada de un usuario físico.
- **Sistema**
 - Debe almacenar los usuarios y las contraseñas en la base de datos.
 - Debe tener un usuario root que permita la gestión de los usuarios.
 - No podrá almacenar en la base de datos usuarios con campos obligatorios nulos.

Requisitos no funcionales

- Cada petición del usuario no debe tardar más de 10 segundos.
- El campo “usuario” y “contraseña” deben ser obligatorios.
- El sistema estará en la capacidad de almacenar los datos de usuarios y contraseña.

Inversos

- El campo de usuario y contraseña no puede ser nulo.
- El campo de entradas no puede ser nulo.

Diagrama de casos de uso

En el siguiente diagrama se encuentran contenidos los elementos del modelo propuesto, los actores y los casos de uso. Aquí, se muestran las diferentes relaciones entre los elementos y sus respectivas acciones.

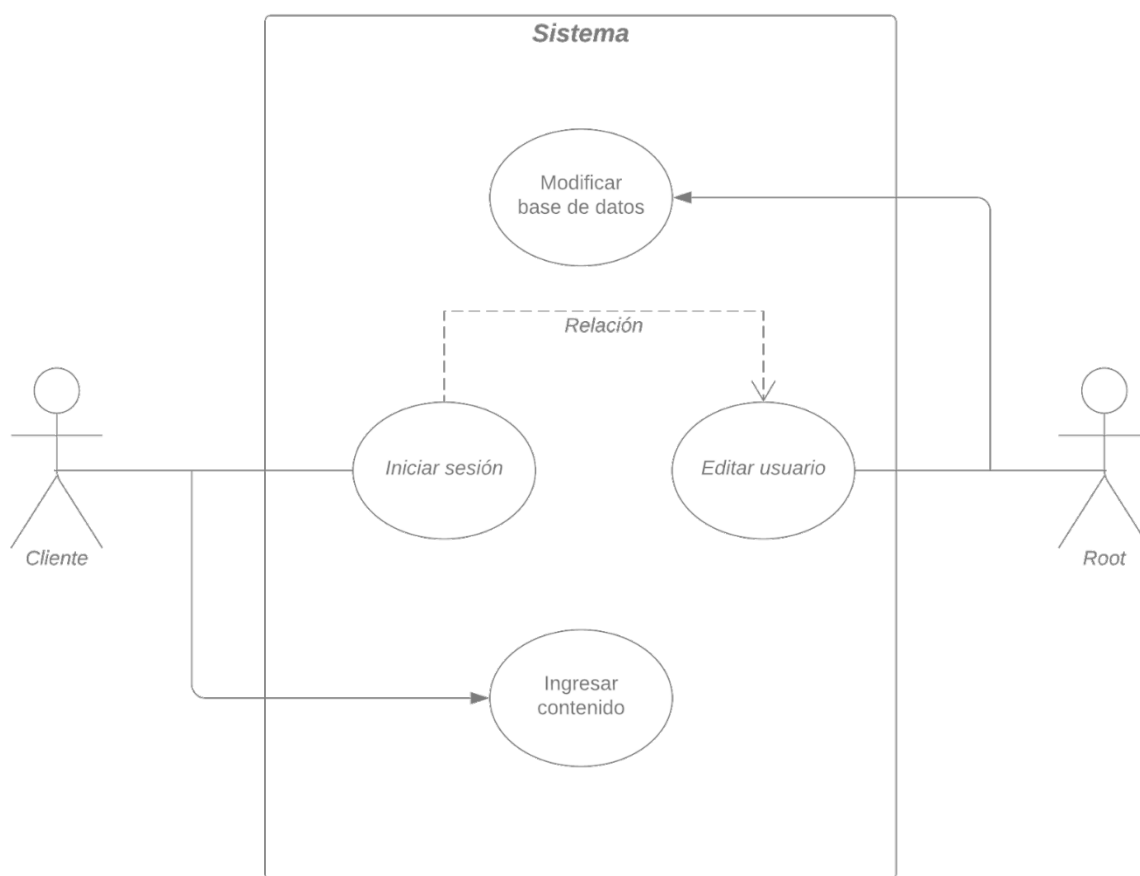


Ilustración 2: Casos de uso. fuente: propia

Diagramas de estados

En los siguientes esquemas planteados para los diagramas de estado se muestran las posibles transiciones entre estados que puede tomar el sistema. El estado de un componente representa un comportamiento observable externamente y que se prolonga limitadamente en el tiempo.

En el proyecto, se proponen inicialmente 5 estados:

Solicitando información para registrarse: En este estado inicial, el profesional de la salud solicita al ente encargado de registrar su ingreso, añadiendo su información a la base de datos del sistema.

Llenando la información solicitada: Aquí, se llenan los datos correspondientes para llevar a cabo el registro del profesional a cargo de usar el sistema, se ingresará su correo corporativo, tarjeta profesional y se le brindará una contraseña.

Verificando la información suministrada: Se validarán que los datos suministrados al administrador del sistema sean correctos.

Retomando a solicitar información de registro: En el caso en que los datos entregados sean correctos, el profesional podrá hacer uso del sistema, de otra manera el proceso se tomará como finalizado.

- Registro de usuario

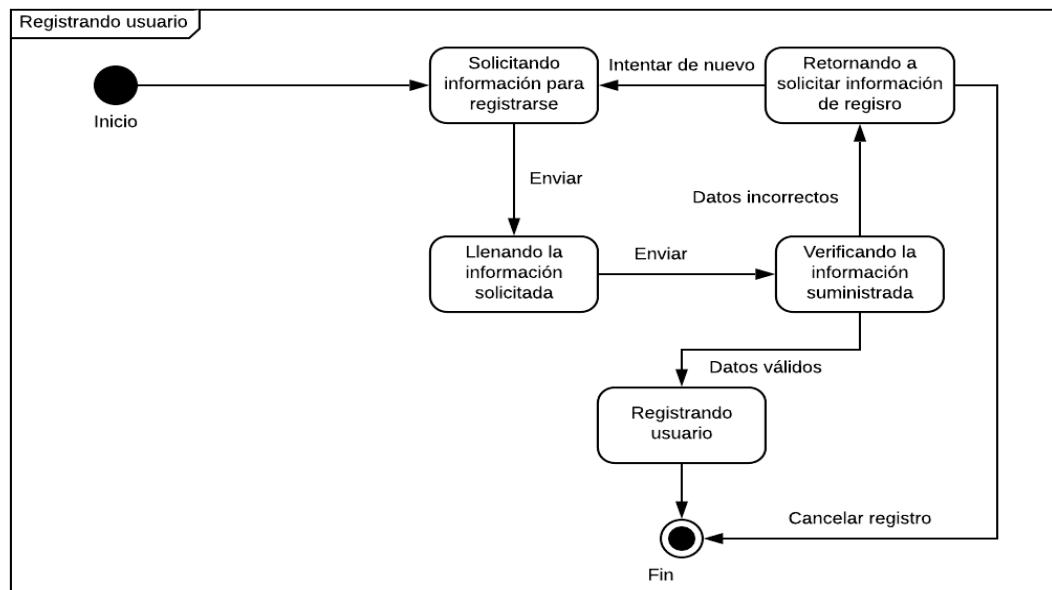


Ilustración 3: DE-Registro de usuarios, fuente: propia

Ingreso datos para iniciar sesión: El encargado de usar el sistema ingresará los datos una vez registrado.

Verificar datos: Los datos suministrados serán comparados con los que se ubican en la base de datos para buscar validarlos.

Ingresando al sistema: Tras recibir los datos validados, se concreta el ingreso al sistema iniciando el proceso convolucional donde la imagen se hace cada vez más pequeña y es separada en patrones de clasificación.

Notificando clave/usuario incorrecto: Se enviará un mensaje de error al recibir y validar que los datos ingresados son incorrectos, volverá a pantalla inicial para volver a intentar.

- Ingreso de usuario

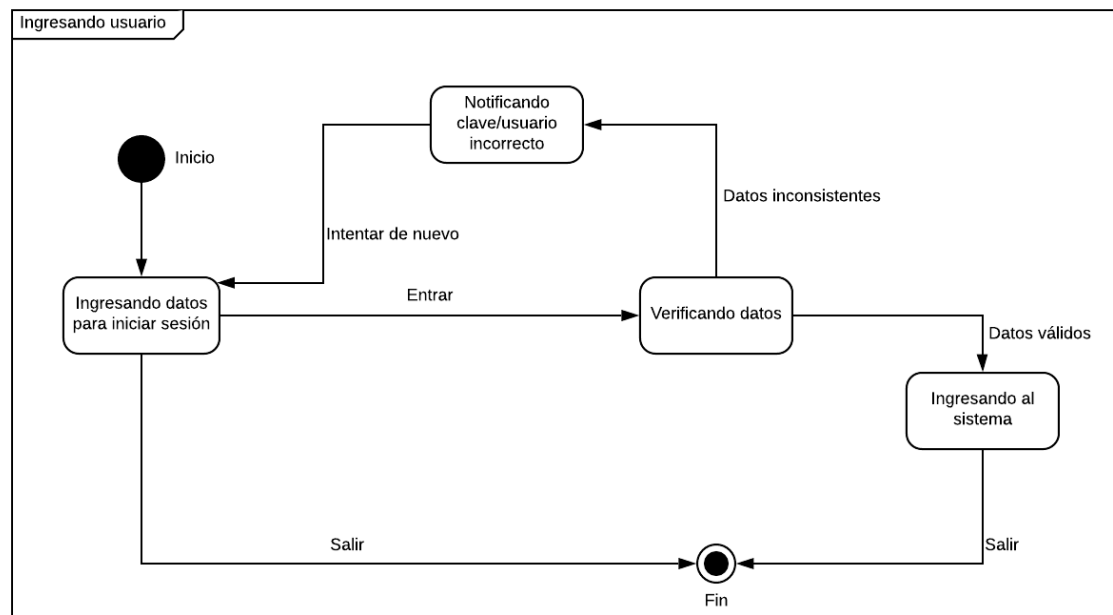


Ilustración 4: DE-Ingresar usuario, fuente: propia

Eligiendo contenido: El profesional elegirá el tipo de imagen que requiere el sistema para empezar el análisis, en este caso el sistema requiere imágenes de tres tipos, axial, sagital y coronal.

Verificando contenido: Se valida que el contenido ingresado coincida con el formato esperado.

Ingresando contenido a evaluar: Una vez el contenido es validado satisfactoriamente, este ingresa al sistema para ser evaluado y termina el proceso.

Notificando contenido erróneo: Si al verificar que el contenido a ingresar no coincide con el formato esperado, podrá volver a intentar subir determinada imagen.

- Ingreso de contenido

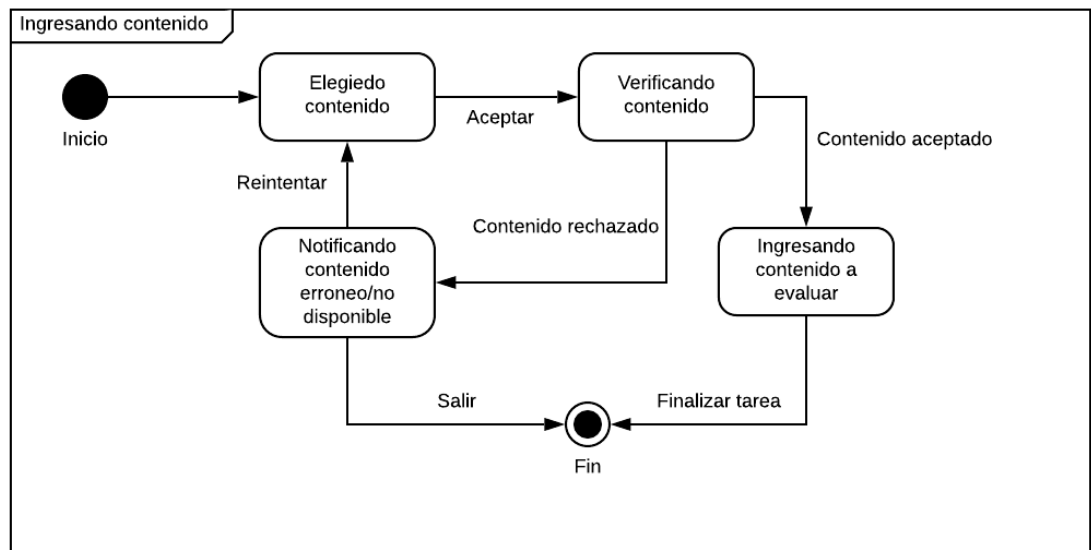


Ilustración 5: DE-Ingresar contenido, fuente: propia

Diagramas de actividades

En los diagramas de actividades, buscamos reflejar el flujo del proceso multipropósito que se usa para modelar el comportamiento del sistema, se puede usar para modelar una clase o un método.

Inicialmente, el usuario ingresa los datos:

Método: El sistema valida en la base de datos y determina si este usuario existe o no

El usuario no existe: El sistema envía un mensaje de error, el usuario puede volver a intentar o finalizar la actividad.

El usuario existe: El sistema permite el ingreso del usuario al sistema y termina la actividad.

- Ingresar usuario

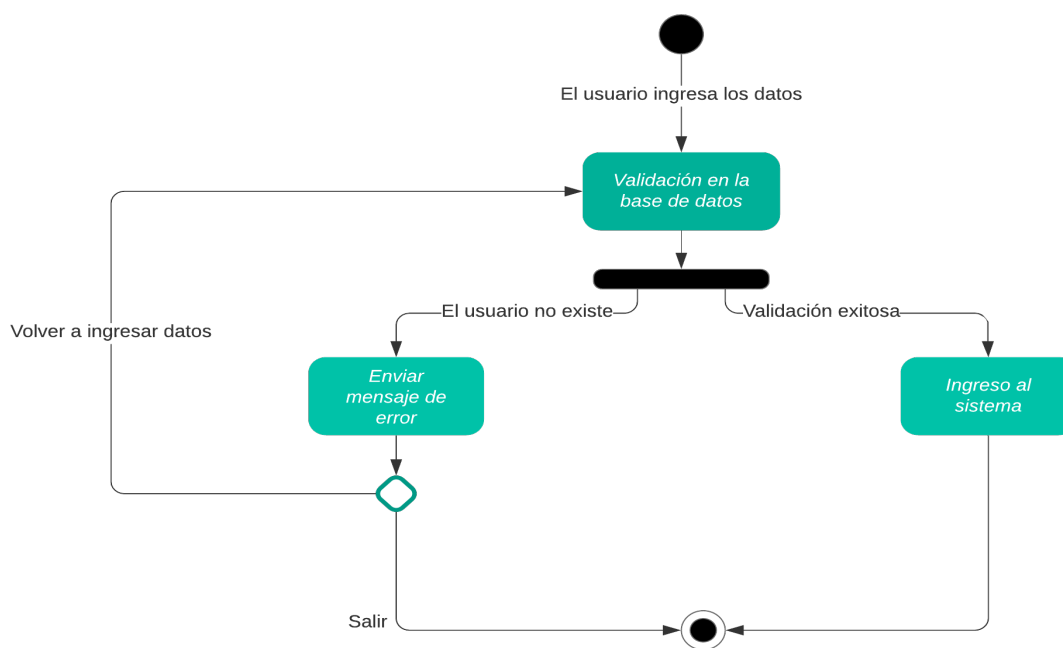


Ilustración 6: DA-Ingresar usuario, fuente: propia

Inicialmente el sistema recibe como entrada en esta actividad que el usuario desea registrarse

Método: El sistema envía el formulario de registro al usuario

El usuario ingresa la información que puede ser correcta o incorrecta

Método: En el caso de que la información se valide correctamente el sistema realiza el registro y la actividad termina, en el caso que la información no sea correcta, el sistema le permitirá al usuario volver a intentar enviando la información o terminar la actividad de registro.

- Registro usuario

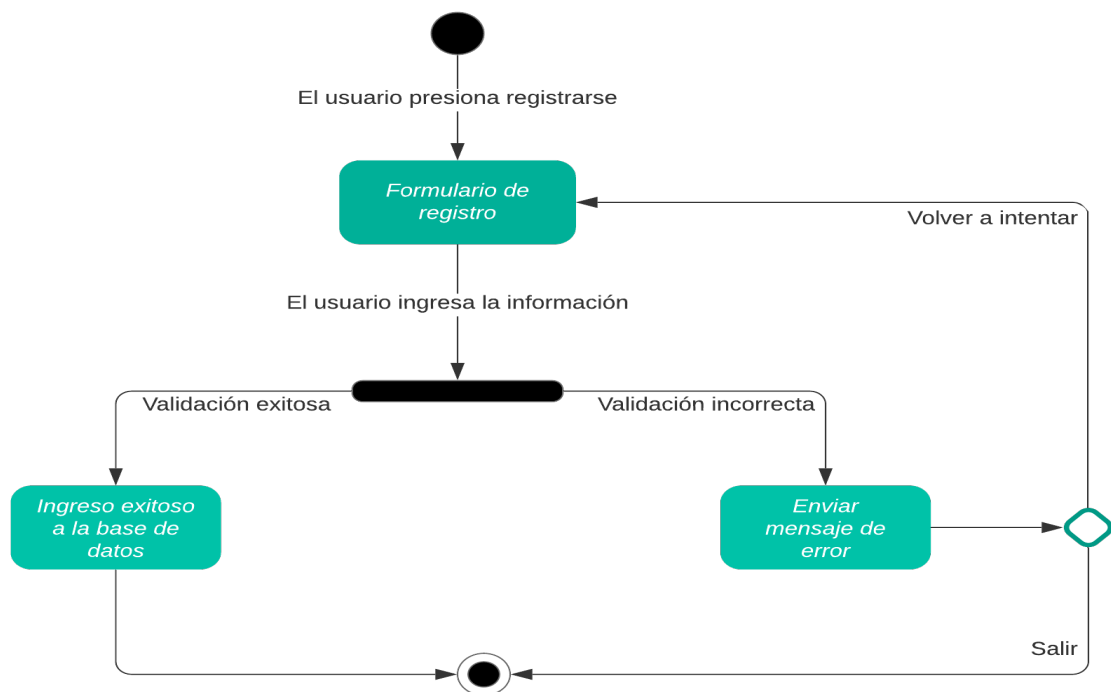


Ilustración 7: DA-Registrar usuario, fuente: propia

- Ingreso contenido

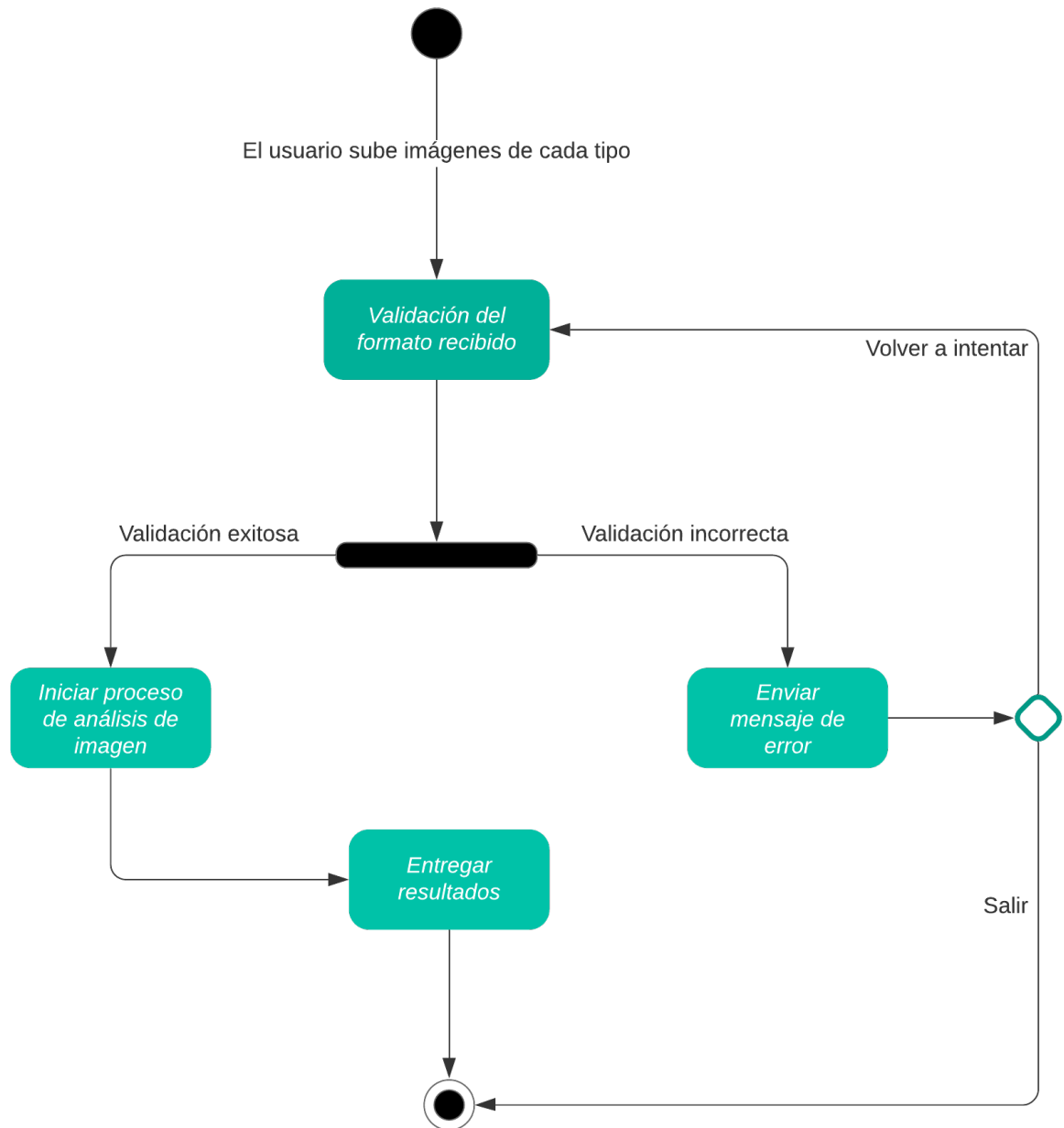


Ilustración 8: DA-Ingresar contenido, fuente: propia

Diagramas de secuencias

En los siguientes diagramas de secuencias se describen básicamente cómo los objetos intercambian mensajes en un orden determinado. Los objetivos son los bloques de construcción y tienen un ciclo de vida limitado y ordenado.

El usuario solicita iniciar sesión a la aplicación y esta la muestra un mensaje de validación o error

La aplicación recibe la solicitud del usuario y solicita verificar la información con la base de datos, finalmente recibe una respuesta con el estado de los datos y envía el mensaje al usuario

La base de datos recibe una solicitud de validación de ingreso desde la aplicación y envía el resultado producto de la validación de los datos comparándolo con su información

- Ingresar usuario

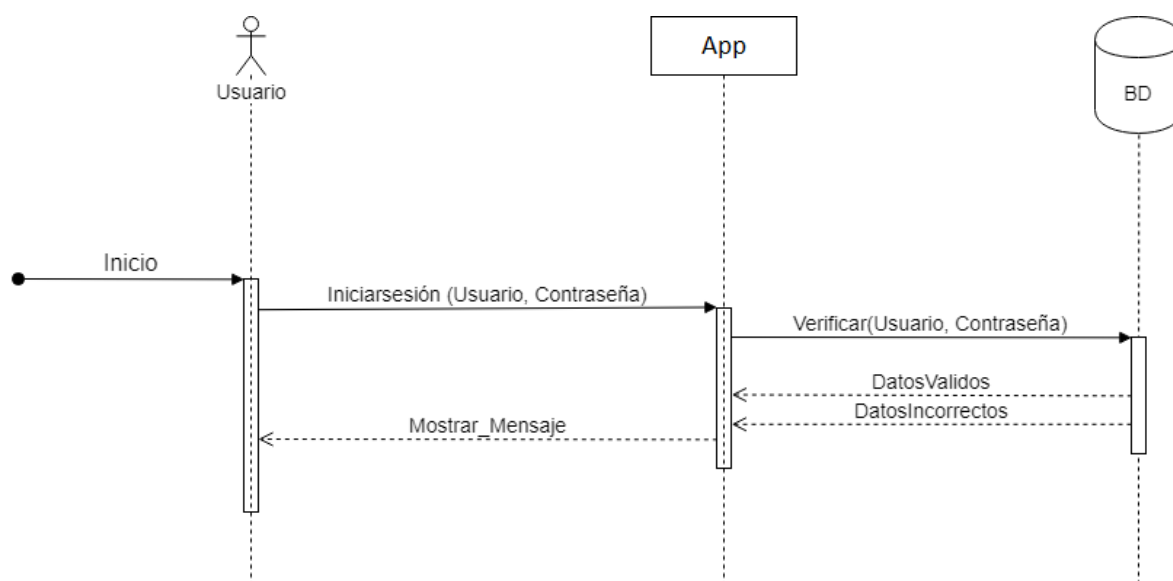


Ilustración 9: DS-Ingresar usuario, fuente: propia

El usuario solicita iniciar sesión a la aplicación y esta la muestra un mensaje de validación o error

La aplicación recibe la solicitud del usuario y solicita verificar la información con la base de datos, finalmente recibe una respuesta con el estado de los datos y envía el mensaje al usuario.

La base de datos recibe una solicitud de validación de ingreso desde la aplicación y envía el resultado producto de la validación de los datos comparándolo con su información

El usuario confirma su registro en la aplicación y recibe una respuesta

La aplicación recibe la confirmación de registro del usuario y la envía a validación con la base de datos, esta misma le responde con una confirmación exitosa de guardado y envía esta información al usuario.

- Registro de usuario

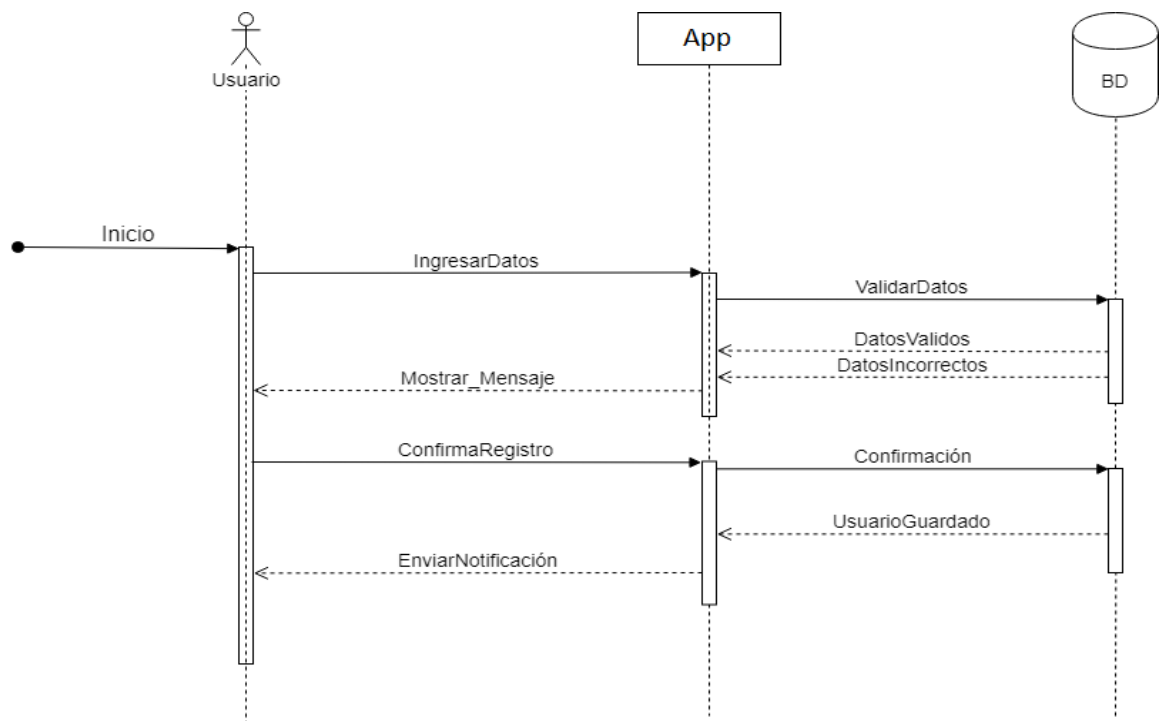


Ilustración 10: DS-Registrar usuario, fuente: propia

El usuario envía contenido a la aplicación y recibe un mensaje luego de haberlos ingresado

La aplicación recibe el mensaje del usuario y los envía a validación en la base de datos, finalmente envía una notificación acerca del contenido ingresado.

La base de datos recibe el contenido, lo valida y envía un resultado a la aplicación.

- Ingreso contenido

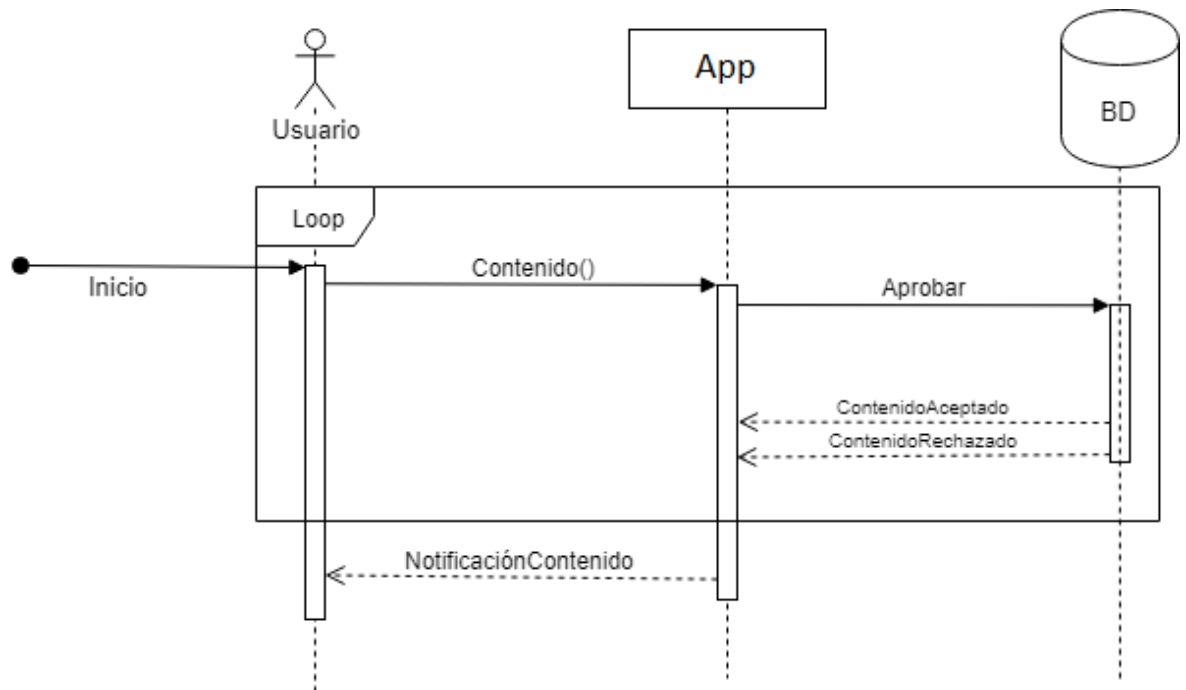


Ilustración 11: DS-Ingresar contenido, fuente: propia

Diagramas de paquetes

En este diagrama de paquetes, se representan las dependencias entre los paquetes que componen el modelo, de manera que muestra cómo el sistema está dividido en agrupaciones lógicas y sus dependencias en este proyecto.

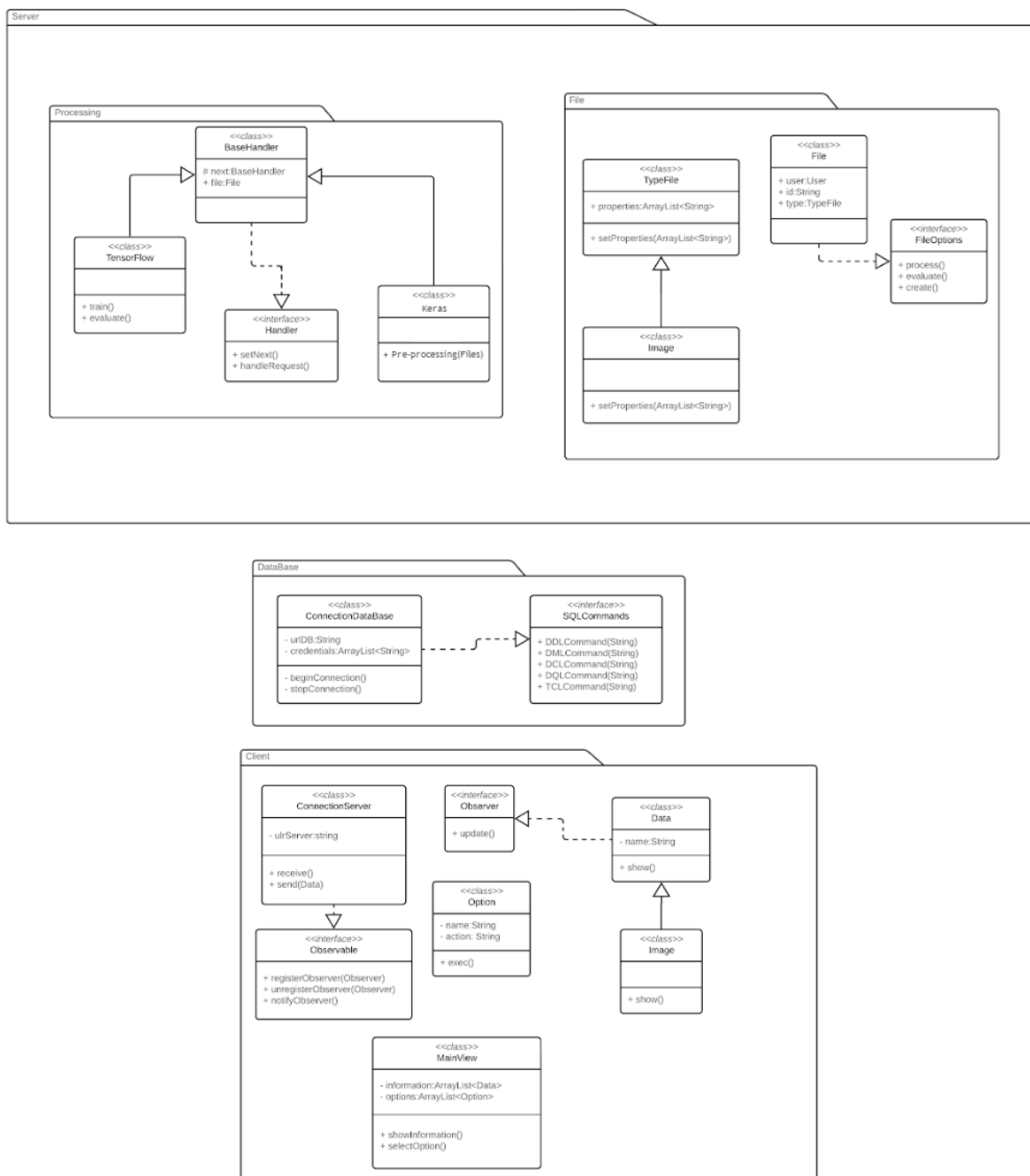


Ilustración 12: Diagrama de paquetes, fuente: propia

Desarrollo y resultados

En la fase uno de la metodología, se recopila la información acerca del cáncer de próstata, con lo cual se evidencia una alta letalidad en personas mayores a 70 años, además, suele ser hereditaria y su avance puede ser tanto agresivo como lento, por lo cual, se dificulta diagnosticar en el momento adecuado. Se halló software que puede ser de apoyo para el profesional de la salud, pero para este cáncer en específico, se utilizan principalmente otros recursos además de las imágenes por resonancia magnética, por ejemplo, las tomas de sangre para medir los niveles de PSA, el cual suele tener un valor elevado en pacientes con esta afección.

Al implementar la imagen por resonancia magnética en la próstata, se intenta generar un programa que, con base en dichas imágenes, esté en capacidad de clasificarlas y que permita evaluar unas nuevas, dándole al paciente un resultado de que posiblemente tenga o no la enfermedad.

Al discutir con el técnico radiólogo David Delgado Buenrostro del gabinete radiológico Imaxess, que desde México quien aclaraba que las imágenes que se toman, no se enfoca directamente a la próstata, se usan tres perspectivas: axial, coronal y sagital, las cuales arrojan diferentes marcos de referencia que el profesional encargado evaluará. Tomando en cuenta la información suministrada por el técnico radiólogo se entiende que, es necesario una red neuronal convolucional por cada corte, ya que los patrones van a diferir de la perspectiva que tengan.

D. Delgado Buenrostro (comunicación personal, 25 de mayo de 2020)

Se llegó a la conclusión de utilizar un banco de imágenes públicas dado que existe una limitación en el acceso a la información de los pacientes, por lo cual, el prototipo no podrá cumplir del todo su propósito, ya que es necesario el acceso a la información de pacientes sin cáncer.

- La base de datos fue dotada con 163 imágenes entre los tres cortes tomados a la próstata.
- El formato de las imágenes es png.

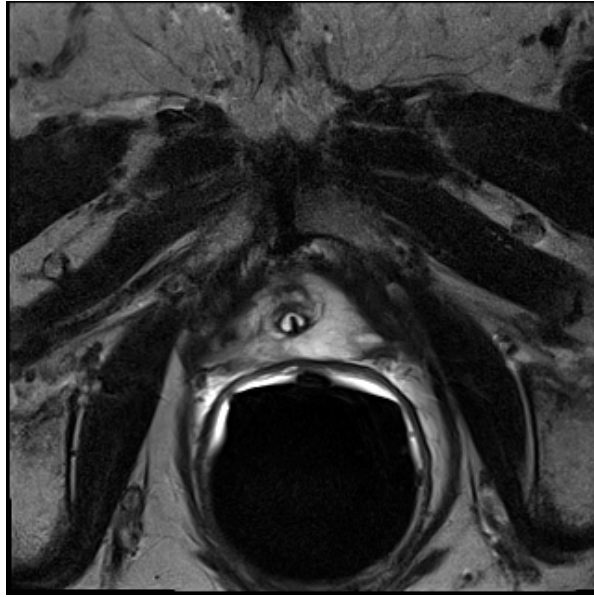


Ilustración 13: Imagen de resonancia magnética, toma axial, fuente: Cancer imaging archive



Ilustración 14: Imagen de resonancia magnética, toma coronal, fuente: Cancer imaging archive



Ilustración 15: Imagen de resonancia magnética, toma sagital, fuente: Cancer imaging archive

Continuando en la fase dos, se utiliza ATAM para tener claro qué módulos del programa serán los más críticos para así priorizar lo más relevante en el la fase de desarrollo, sumando los diagramas en UML para visualizar gráficamente las funcionalidades y roles que se cumplirán en el sistema.

En la fase tres, para la interfaz de usuario se usó html5 css3 y, el lenguaje elegido para el entrenamiento de la red neuronal convolucional fue Python ya que, las librerías TensorFlow 2.2.0 y Keras 2.4.3 que posee las funciones para redes neuronales usada en el aprendizaje automático son complemento de este lenguaje, además, ambas bibliotecas son de código abierto y, finalmente la base de datos es una separación por carpetas en la cual cada una tendrá su respectivo corte y las imágenes dispuestas para cada clase, con las cuales entrenamos la red.

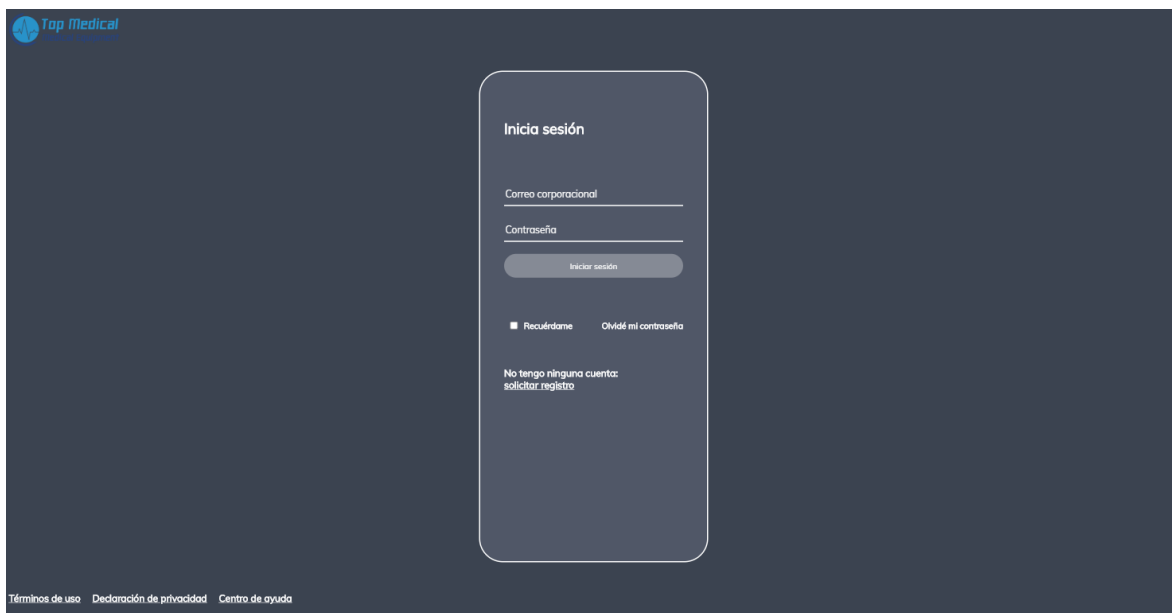


Ilustración 16: Vista de inicio de sesión, fuente: propia

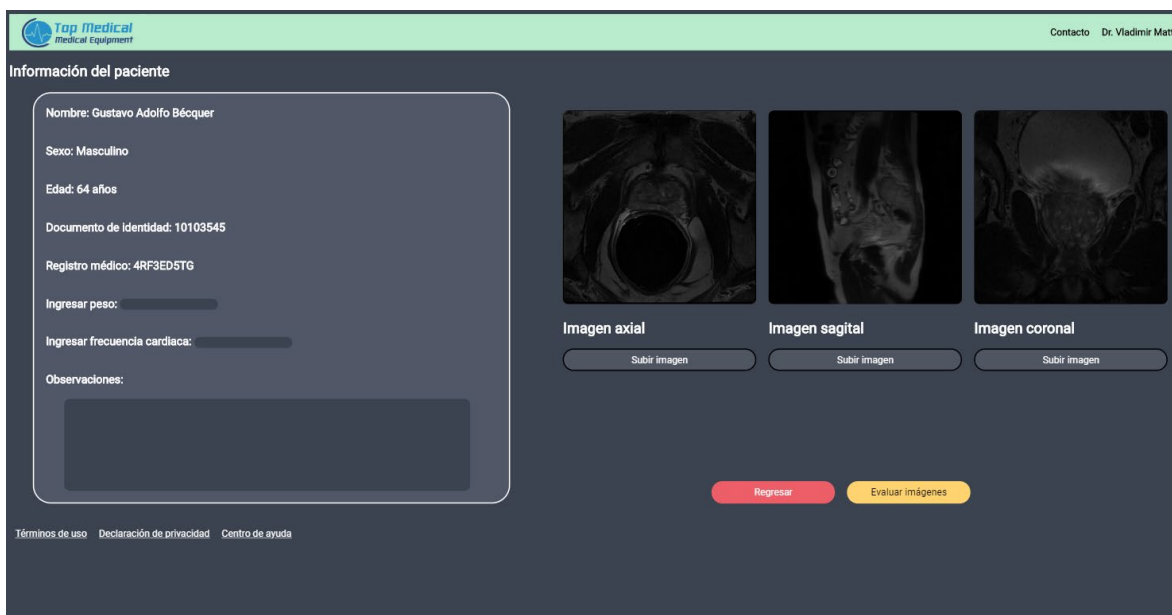


Ilustración 17: Vista ingreso datos, fuente: propia

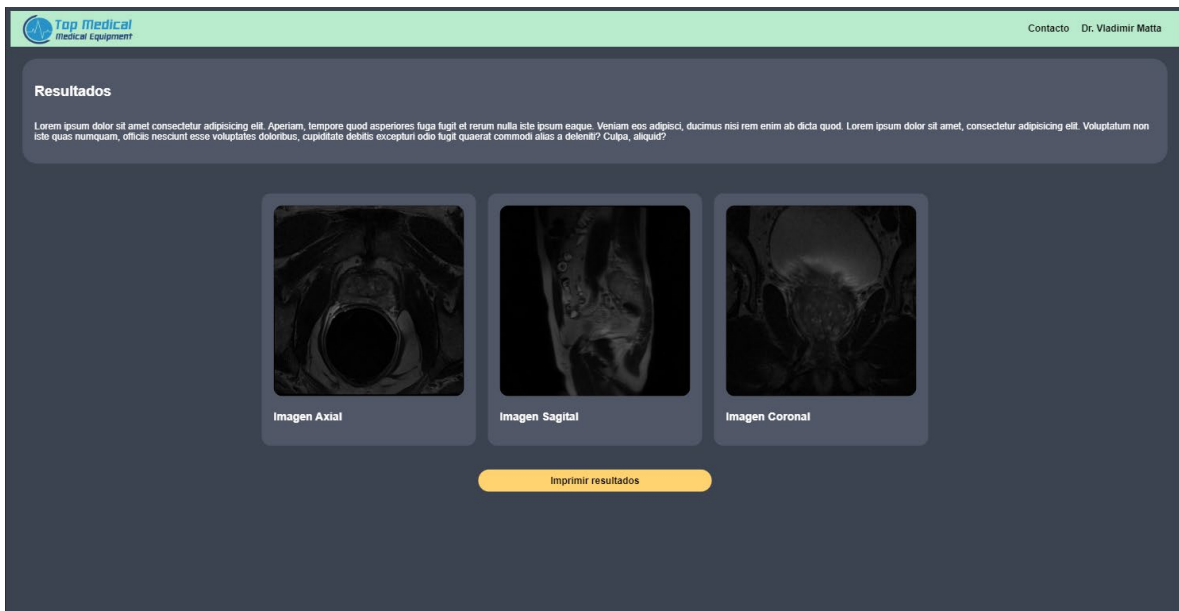


Ilustración 18: Vista resultado evaluación, fuente: propia

```

Parameters
epoch=20
height, width = 8, 8
batch_size=32
steps = 1000
validation_steps=300
filtersConv1=32
filtersConv2=64
size_filter1=(2, 2)
size_filter2=(1, 1)
size_pool=(1, 1)
classes=2
lr=0.0004

```

Ilustración 19: Parámetros para la red neuronal, fuente: Propia

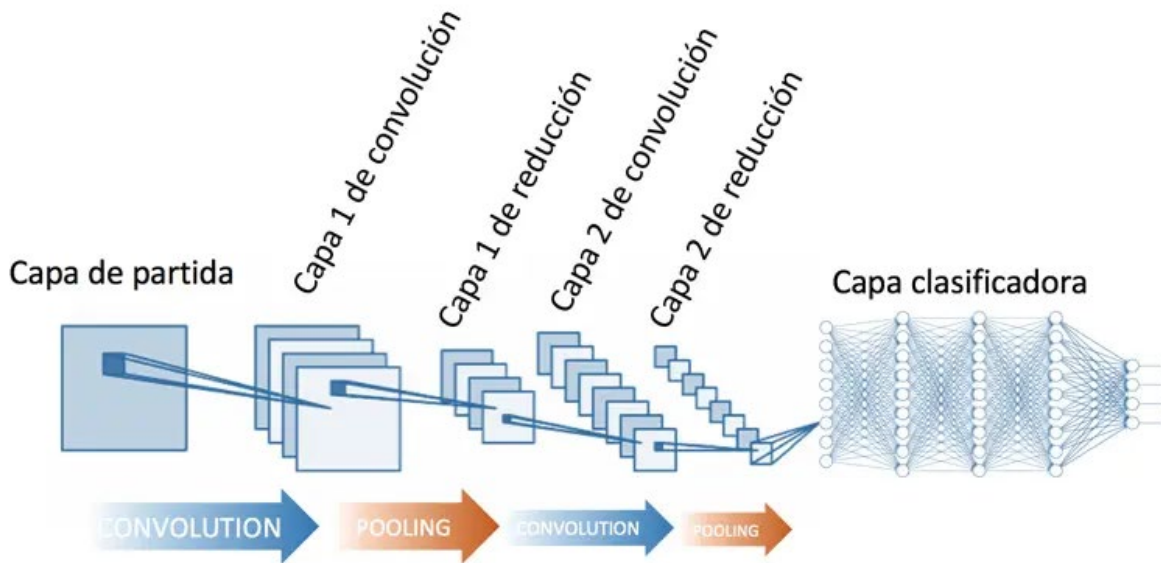


Ilustración 20: Esquema red neuronal convolucional, fuente: <https://www.diegocalvo.es/>

Se implementan tres redes con los mismos parámetros que constan de nueve capas, antes de agregar las capas se usa la función *Sequential* la cual irá ejecutando capa por capa en orden descendente, seguido a esto se ejecutan dos ramas idénticas, donde se ejecuta *MaxPooling2D* el cual detectará el número más alto del parche que seleccionará y los pasará a una matriz más pequeña y *Convolution2D* que detectará los patrones en la imagen, y en cada paso del entrenamiento irá generando imágenes más pequeñas de salida; cabe aclarar que “size_filter” tendrá un valor diferente en cada rama para que evalúe de manera diferente.

```
cnn = Sequential()
cnn.add(Convolution2D(filtersConv1, size_filter1, padding = "same", input_shape=(height, width, 3), activation='relu'))
cnn.add(MaxPooling2D(pool_size=size_pool))

cnn.add(Convolution2D(filtersConv2, size_filter2, padding = "same"))
cnn.add(MaxPooling2D(pool_size=size_pool))
```

Ilustración 21: Capas de inicio, convolución y agrupación, fuente: propia

La capa *Flatten* hará que la imagen ahora solo posea una dimensión que tiene toda la información sobre esta.

```
cnn.add(Flatten())
```

Ilustración 22: Capa aplanar, fuente: propia

En la capa *Dense* se especifica una cantidad de neuronas (512) que irán conectadas a la capa de entrada y la función de activación es *relu* ya que los valores que sean negativos de la imagen no son útiles.

```
cnn.add(Dense(512, activation='relu'))
```

Ilustración 23: Capa densa, fuente: propia

La capa *Dropout* apaga aleatoriamente las neuronas que se generan en la capa *Dense* para que el sistema no se sature y genere patrones no deseados.

```
cnn.add(Dropout(0.5))
```

Ilustración 24: Capas de expulsión, fuente: propia

otra capa *Dense*, esta vez la cantidad de neuronas serán igual a la cantidad de clases y la función de activación 'softmax' dará un porcentaje a cada clase, que al sumar, su resultado dará 1,

```
cnn.add(Dense(classes, activation='softmax'))
```

Ilustración 25: Capa densa para elección de clase, fuente: propia

finalmente, la capa *Compile* que se compone de tres funciones, la función de pérdida 'binary_crossentropy', elegida por solo tratarse de dos clases, el optimizador usado es Adam, que maneja actualizaciones dispersas de manera más eficiente.

```
cnn.compile(loss='binary_crossentropy',  
            optimizer=optimizers.Adam(lr=lr),  
            metrics=['categorical_accuracy'])
```

Ilustración 26: Capas de salida, fuente: propia

Se prueba individualmente cada red con su corte asignado, el entrenamiento se ejecutó en un computador con las siguientes características: sistema operativo Windows 10 Pro x64 con procesador FX 8320 de 8 núcleos a 3.5 Ghz y 16 Gb de RAM a 1666 Mhz; el tiempo empleado para su entrenamiento fue de cuatro horas para cada red, ejecutándose desde el Kernel del sistema, la cual genera los archivos 'modelo.h5' y 'pesos.h5' por cada red, que permitirán guardar los datos y no sea necesario entrenarse cada vez que se inicie el programa.



 modelo.h5	Archivo H5	393.292 KB
 pesos.h5	Archivo H5	131.106 KB

Ilustración 27: Evidencia de red entrenada, fuente: propia

y por último, se crea una función que reciba la información de las tres redes para evaluar las tres imágenes ingresadas por el usuario en simultáneo y retorne una respuesta única.

```
def predict(file1, file2, file3):
    x = load_img(file1, target_size=(height, width))
    x = img_to_array(x)
    x = np.expand_dims(x, axis=0)
    array = cnn.predict(x)
    result1 = array[0]
    answer1 = np.argmax(result1)
    y = load_img(file1, target_size=(height, width))
    y = img_to_array(y)
    y = np.expand_dims(y, axis=0)
```

Ilustración 28: Fragmento función de salida del programa, fuente: propia

Conclusiones

- En el proceso se determinó que la luminosidad no sería un factor tenido en cuenta, ya que con anterioridad se eligió que el parámetro tenido en cuenta en este proyecto serían patrones de formas que detectara la red.
- Se destaca que las librerías TensorFlow y Keras, que son de código abierto, son una gran herramienta al usar Python, dado que las mismas están hechas bajo ese lenguaje, de manera que se hace más simple para el desarrollo.
- El uso de herramientas tecnológicas con base en inteligencia artificial como insumo al chequeo médico resulta útil, sin embargo, no sustituye la metodología tradicional porque el resultado se determina por lo que arroja la biopsia.

D. Delgado Buenrostro (comunicación personal, 25 de mayo de 2020)

- Es necesario evaluar todos los cortes/ejes dado que es parte del procedimiento realizado para diagnosticar cáncer de próstata.

D. Delgado Buenrostro (comunicación personal, 25 de mayo de 2020)

Aportes

- Profundizar en las imágenes de resonancia magnética bajo la perspectiva de sistemas computacionales para que sirva como herramienta de diagnóstico y ayude al profesional a la detección temprana del cáncer de próstata sin usar métodos invasivos.

Recomendaciones

- TensorFlow puede hacer uso de las GPU de Nvidia, lo cual acelera el proceso de entrenamiento de la red neuronal convolucional.
- Se propone un aumento en la complejidad de la red neuronal convolucional para precisar más la clasificación de las imágenes, ya que los detalles a evaluar pueden ser de un tamaño bastante reducido.
- Si se usa un formato diferente al dcm, se recomienda el uso del formato jpeg en la conversión para perder la menor cantidad de información en los píxeles.
- Se recomienda una resolución de imagen de por lo menos 160 ppp (píxeles por pulgada) para aumentar precisión.
- Es posible tomar como base este prototipo para futuros desarrollos de diagnóstico añadiendo mayor complejidad desde el ámbito médico: con el cual sea posible dar un resultado con mayor fiabilidad, y desde el ámbito informático: agregando la lógica difusa para resultados más cercanos a la realidad.

Bibliografía

- [1] *Nanopharmacia, herramienta tecnológica Aurora (2019).*
<https://elmedicointeractivo.com/aurora-el-futuro-en-actualizacion-inmediata-para-oncologos-y-genetistas/>
- [2] *The 4Kscore® Test: A blood test for Assessing the Risk of aggressive prostate cancer.*
<https://www.4kscore.com>
- [3] *Kang, S., Li, Q., Chen, Q. et al. CancerLocator: non-invasive cancer diagnosis and tissue-of-origin prediction using methylation profiles of cell-free DNA. Genome Biol 18, 53 (2017).* <https://doi.org/10.1186/s13059-017-1191-5>
- [4] *Guía práctica clínica (GPC) para la detección temprana, diagnóstico, tratamiento, seguimiento y rehabilitación del cáncer de próstata (2013)*
<https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/INEC/IETS/Cancer-prostata-fina-Prof-Salud.pdf>
- [5] *Redacción salud. El espectador (2017).*
<https://www.elespectador.com/noticias/salud/cancer-de-prostata-un-estudio-inquietante-articulo-703343>
- [6] *Magnetic Resonance Imaging (MRI) – Prostate (Revisión febrero 2020)*
https://www.radiologyinfo.org/en/info.cfm?pg=mr_prostate
- [7] *Wallis CJ, Haider MA, Nam RK. Role of mpMRI of the prostate in screening for prostate cancer. Transl Androl Urol 2017;6(3):464-471. doi: 10.21037/tau.2017.04.31*
- [8] *García-Perdomo HA, Zapata-Copete JA, Sánchez A. Una mirada global y actualizada del cáncer de próstata. Rev. Fac. Med. 2018;66(3):429-37. Spanish. doi: http://dx.doi.org/10.15446/revfacmed.v66n3.65770*
- [9] *Pereira PP, Ramírez Vargas RD, y Fernández Calvo O. (última actualización 2018). Cáncer de próstata. Elsevier. Disponible en: https://www-clinicalkey-es.ezproxy.utp.edu.co/#!/content/guides/techniques/52-s2.0-mt_fis_52*
- [10] *Ruiz López AI, Pérez Mesa JC, y Cruz Batista Y, González Lorenzo LE. (2015). Actualización sobre el cáncer de próstata. Correo científico médico de Holguín, CCM 2017;(3) 876-887. Disponible en: http://scielo.sld.cu/pdf/ccm/v21n3/ccm21317.pdf*
- [11] *Rivero, A., y Berríos, R. (2016). El cáncer de próstata y la construcción social de la masculinidad en Puerto Rico. Psicología, Conocimiento y Sociedad, 6(1), 164-190. http://www.scielo.edu.uy/scielo.php?script=sci_arttext&pid=S1688-70262016000100008*

[12] Villegas CA, Chacón JA, Sánchez T. (2015). Sobrevida en cáncer de próstata de una población del centro de Colombia. Disponible en:

<http://www.scielo.org.co/pdf/amc/v40n2/v40n2a06.pdf>

[13] Guía práctica clínica (GPC) sobre el proyecto clínico que adelanta BioMedIA (2017)

<https://biomedia.doc.ic.ac.uk/project/ifind/>

[14] Cancer Incidence and mortality worldwide

<https://onlinelibrary.wiley.com/doi/full/10.1002/ijc.29210>

[15] Boletín Epidemiológico INC Colombia ISSN 2149 - 0196

<https://www.cancer.gov.co/files/libros/archivos/Boletin>